replaces California Institute of Technology Social Science Working Paper 123

*Econometrica*, Vol. 45, No. 7 (October, 1977)

# A NOTE ON DISTRIBUTED LAGS, PREDICTION, AND SIGNAL EXTRACTION

BY DAVID M. GRETHER

## 1. INTRODUCTION

A WIDE VARIETY OF ECONOMIC models include as explanatory variables either expectational variables or variables representing the result of some decision-making process. The first category includes both expectations about the future values of variables, e.g., next period's sales, the level of unemployment two quarters ahead, etc. and other subjective variables such as permanent income or the "normal" level of prices and interest rates. Examples of the second type are "desired" capital stock, planned production, or inventory accumulation, and so on.

Since data on expectations or specific decisions are frequently unavailable, these models are often made empirically testable by specifying the way in which the expectational or choice variables are related to observable quantities. As is well known, this specification often leads to a distributed lag model such as the familiar adaptive expectations model or the stock adjustment model. In this note we consider models of the form

$$y_t = a + bx_t^* + u_t \tag{1}$$

where $x_t^*$ is one of the types of variables mentioned above. We restrict our attention to situations in which the exogenous variables have mixed autoregressive, moving average representation (i.e., have rational spectral densities) and in which $x_t^*$ is chosen to minimize the expected value of a quadratic objective function. For example, $x_t^*$ could be the least squares forecast of $x_{t+j}$ calculated at time $t$.

It turns out, not surprisingly, that in these cases the resulting distributed lag model is a rational distributed lag [**6**]. More importantly, the orders of the polynomials in the lag operator depend in a simple way upon the stochastic structure of the exogenous variables and upon the nature of the optimization problem. The results reported here should be of practical use as *ARMA*-type specifications are quite common as is the use of quadratic objective functions. The results are not, strictly speaking, new in that they are implicit in Whittle's work [**10**], and special cases have been worked out by Nerlove [**7**], though no general exposition exists in the econometric literature.

In what follows it is assumed that all stochastic processes are zero mean covariance stationary processes with autoregressive representation.[1] We use the following notational conventions:

(i) $g_{yx}(z) = \sum_{k=-\infty}^{\infty} E(y_t x_{t-k}) z^k.$

(ii) If $H(z) = \Sigma_{-\infty}^{\infty} h_i z^i$ is the Laurent expansion of a function which converges in an annulus containing the unit circle, then $[H(z)]_+ = \Sigma_0^{\infty} h_i z_i$, and $[H(z)]_- = \Sigma_{-\infty}^{-1} h_i z_i$.

(iii) $L^k \cdot x_t \equiv x_{t-k}$.

[1] The results presented can be extended to processes whose $p$th differences are as stated, but the less general case is assumed for ease of exposition. See [**10**, Ch. 8, esp. pp. 92–96]. Also, without the additional assumption that the processes are Gaussian, the solutions presented need to be interpreted as optimal only in the class of linear rules.

## 2. SIGNAL EXTRACTION AND PREDICTION

Let

$$(2)\qquad w_t = x_t + \eta_t,$$

$$x_t = \frac{N(L)}{D(L)}\mathscr{E}_t,$$

$$\eta_t = \frac{R(L)}{S(L)}\zeta_t,$$

where $N(\ )$, $D(\ )$, $R(\ )$, and $S(\ )$ are polynomials of degree $n$, $d$, $r$, and $s$, respectively, and $\{\mathscr{E}_t\}$, $\{\zeta_t\}$ are mutually uncorrelated white noise sequences.[2] Further, let $T(\ )$ be a polynomial of degree $\max(n+s, d+r)$ satisfying:

$$(3)\qquad \sigma^2 T(z)T(z^{-1}) = \sigma_{\mathscr{E}}^2 N(z)S(z)N(z^{-1})S(z^{-1}) + \sigma_\zeta^2 R(z)R(z^{-1})D(z^{-1})D(z)$$

with the roots of $T(\ )$ lying outside the unit circle and $\sigma^2$ chosen so that $t_0 = 1$.

THEOREM (SIGNAL EXTRACTION): *Let $\hat{x}_{t+v,t}$ be the least equares estimate of $x_{t+v}$ made at time t based upon observations on $w_s$, $s \leq t$. Then*

$$\hat{x}_{t+v,t} = \gamma(L)w_t,$$

$$\gamma(z) = \frac{S(z)N_v(z)}{T(z)},$$

*where $N_v(\ )$ is a polynomial of order* $\max(n-v, d-1, 0)$.[3]

PROOF: From (2) we have

$$g_{wx}(z) = g_{xx}(z) = \sigma_{\mathscr{E}}^2 \frac{N(z)N(z^{-1})}{D(z)D(z^{-1})},$$

$$g_{ww}(z) = \sigma^2 \frac{T(z)T(z^{-1})}{D(z)S(z)D(z^{-1})S(z^{-1})}.$$

Thus by [**10**, p. 42, eq. 2]

$$(4)\qquad \gamma(z) = \frac{D(z)S(z)}{\sigma^2 T(z)}\left[\frac{\sigma_{\mathscr{E}}^2 N(z)N(z^{-1})S(z^{-1})D(z^{-1})}{D(z)D(z^{-1})T(z^{-1})z^v}\right]_+$$

$$= \frac{D(z)S(z)}{\sigma^2 T(z)}\left[\frac{\sigma_{\mathscr{E}}^2 N(z)N(z^{-1})S(z^{-1})}{T(z^{-1})D(z)z^v}\right]_+ .$$ [4]

[2] As we are assuming that all processes have autoregressive representations, these polynomials are assumed to have all their zeros outside the unit circle. Clearly, for identification we must assume that certain pairs of polynomials, e.g., $N(\ )$ and $D(\ )$ have no common zeros. To keep the exposition as smooth as possible, these types of assumptions are implicit throughout.

[3] Thomas McCoy has pointed out that certain kinds of coefficient restrictions can reduce the order of $N_v(\ )$. Without a priori knowledge of such restrictions one would have to allow for lags of the order indicated.

[4] [**10**, equation 3.7.2] gives $\gamma(z) = [1/B(z)][g_{vx}(z)/B(z^{-1})]_+$ where $g_{xx}(z) = \sigma^2 B(z)B(z^{-1})$ and $B(z)$ has all its zeros outside the unit.

The expression under the $[\ ]_+$ operator can be evaluated using Theorem 1 on page 93 of [**10**]. To see this, note that

$$\left[\frac{\sigma_{\mathscr{E}}^2 N(z)N(z^{-1})S(z^{-1})}{T(z^{-1})D(z)z^v}\right]_+ = \frac{\left[\dfrac{\sigma_{\mathscr{E}}^2 N(z)N(z^{-1})S(z^{-1})}{T(z^{-1})z^v}\right]_+}{D(z)}$$
$$+\left[\frac{\left[\dfrac{\sigma_{\mathscr{E}}^2 N(z)N(z^{-1})S(z^{-1})}{T(z^{-1})z^v}\right]_-}{D(z)}\right]_+ .$$

The first term is clearly of the form $A_1(z)/D(z)$ where $A_1(\ )$ is of order max $(n-v, 0)$. To obtain the second term we may expand $1/D(z)$ by partial fractions and apply Whittle's theorem to each term in the resulting sum. On recombining terms, the second expression is of the form $A_2(z)/D(z)$ where $A_2(\ )$ is of order $d-1$. *Q.E.D.*

COROLLARY (Prediction): *Let $x_t = (N(L)/D(L))\mathscr{E}_t$ where $\{\mathscr{E}_t\}$ is white noise and $N(\ )$ and $D(\ )$ are polynomials of degree $n$ and $d$ respectively. The least squares forecast of $x_{t+v}$ made at time $t$ based upon observation on $x_s$, $s \leqslant t$ is given by*

$$\hat{x}_{t+v} = \frac{N_v(L)}{N(L)} x_t$$

*where $N_v(\ )$ is a polynomial of degree max $(n-v, d-1, 0)$.*

In this case $g_{xx}(z) = g_{ww}(z)$, and the same proof works with $T(z) \equiv N(z)$, and $S(z) \equiv 1$.

If in the structural model (1), $x_t^*$ is the least squares forecast of some covariance stationary process, either based on past observations on the process itself or upon observations with (serially correlated) measurement error, the model becomes

$$(1)' \qquad y_t = a + b\gamma(L)x_t + u_t.$$

Except for the case of finite order autoregressions observed without error, these forecasts in general depend upon the entire past history of the $x$ series. The preceding result shows that for arbitrary autoregressive, moving average processes, the lag distribution is rational and the orders (or at least upper bounds on them) may be obtained by examining the properties of the observed exogenous variable.

Models employing expectations about future levels of observable economic variables are sufficiently common that citing examples seems unnecessary. For some examples explicitly using expectations about unobserved components of economic time series, see [**7**] and the references there.

## 3. OTHER APPLICATIONS

Suppose that the decision problem is not forecasting or estimating a noise corrupted signal, but, instead, it is to optimize an objective function which depends upon the future values of a time series or upon some unobserved component of a time series. It is well known that if the objective function is suitably restricted, the unknown variables may be replaced by their conditional expectations and the solution obtained in terms of the certainty equivalents [**9**]. Replacing these conditional expectations by the optimum forecasts or extractions will then lead to a distributed lag model. As before the order of the lag operators will depend relatively simply upon the characteristics of the process being forecast and upon the nature of the objective function. While this does provide a generalization of the results of the previous section, we emphasize at the outset that the approach has some severe limitations. First, it is restricted to problems in which the objective function is quadratic which rules out many, perhaps most, interesting applications. Also, the restriction to considering only linear decision rules or normal processes ought to be reemphasized.

Consider the following generalization of the prediction problem:

$$\min_{\gamma} E\{((A(L)\gamma(L)+C(L))x_t)^2\}, \tag{5}$$

$$\gamma(z)=\sum_{i=k}^{\infty}\gamma_i z^i, \qquad g_{xx}(z)=\sigma^2 B(z)B(z^{-1}).$$

The solution is

$$\begin{aligned}\gamma(z) &= -\frac{1}{B(z)A(z)}[B(z)C(z)]_k \\ &= -\frac{z^k}{B(z)A(z)}\left[\frac{B(z)C(z)}{z^k}\right]_+.\end{aligned} \tag{6}$$

[5]

If $B(z)=N(z)/D(z)$, then the corollary above gives the order of

$$\frac{D(z)}{N(z)C(z)}\left[\frac{C(z)N(z)}{D(z)z^k}\right]_+=\frac{N_k(z)}{C(z)N(z)},$$

where the order of $N_k(\ )$ is $\max(c+n-k, d-1, 0)$. Thus one can easily determine the order of $\gamma(\ )$.

If instead of (4) we wish to minimize the sum of several such terms, e.g.,

$$E\left\{\sum_{i=1}^{p}\lambda_i[(A_i(L)\gamma(L)+C_i(L))x_t]^2\right\},$$

then the solution is

$$\gamma(z)=-\frac{z^k}{A(z)B(z)}\left[\frac{B(z)\sum_{i=1}^{p}\lambda_i C_i(z)A_i(z^{-1})}{A(z^{-1})z^k}\right]_+, \tag{7}$$

where $A(z)A(z^{-1})=\sum_{i=1}^{p}\lambda_i A_i(z)A_i(z^{-1})$ and $A(z)$ has its roots outside the unit circle.

As an example consider the case of a firm which produces to stock; i.e., holds inventories. Assume that the firm's costs in period $t$ are given by

$$C_t=\lambda_1(P_t-P_{t-1})^2+\lambda_2(I_t-\alpha S_t)^2 \tag{8}$$

where $P_t$ is production in period $t$, $S_t$ is sales in period $t$, $I_t$ is inventories at the end of period $t$, and

$$P_t=S_t+I_t-I_{t-1}.$$ [6]

We also assume that the firm must choose the level of production for period $t$ before the amount of sales for that period is known. Given the accounting identity between production sales, and the change in inventories, it suffices to determine either inventory holdings or the rate of production. While formally it makes little difference, it seems more natural to assume that it is the rate of production which is decided upon rather than the level of inventories. This makes any discrepancy between expected and actual sales show up as unplanned inventory accumulation rather than in unanticipated fluctuations in the rate of

[5] See [**10**, pp. 118–122, esp. equation 10.5.11].

[6] This problem is similar to, but simpler than, those treated by [**1**, **4**, and **5**]. It is presented as an example for expository purposes only and is not intended to be very realistic. One could allow for deterministic components in the series and add linear terms to the cost function without adding any essential complications. See [**10**, section 10.6].

production. So let

$$(9)\qquad P_t=\gamma(L)S_t=\sum_{i=1}^{\infty}\gamma_i S_{t-i}.$$

The expected costs for period $t$ are given by

$$(10)\qquad E(C_t)=E\{\lambda_1(P_t-P_{t-1})^2+\lambda_2(I_t-\alpha S_t)^2\}.$$

If the firm behaves so as to minimize

$$V=E\left\{\sum_{v=0}^{\infty}\rho^v C_{t+v}\right\},$$

the problem is

$$(11)\qquad \min_{\{P_{t+v}\}} E\left\{\frac{\lambda_1(P_t-P_{t-1})^2+\lambda_2(I_t-\alpha S_t)^2}{1-\rho L^{-1}}\right\}$$

where $P_t=\gamma(L)S_t$.

Consider, first, the case in which the firm acts to minimize costs in period $t$. Assuming that the current level of inventories is simply the sum of past differences between production and sales (i.e., initial inventories $I_0$ equals zero), the problem reduces to

$$\min_{\gamma} E\left\{\lambda_1((1-L)\gamma(L)S_t)^2+\lambda_2\left[\left(\frac{\gamma(L)-1}{1-\delta L}-\alpha\right)S_t\right]^2\right\}|\delta|<1.^7$$

This problem is clearly of the type just discussed with

$$A_1(L)=1-L,$$
$$C_1(L)=0,$$
$$A_2(L)=\frac{1}{1-\delta L},$$
$$C_2(L)=-\frac{1+\alpha(1-\delta L)}{1-\delta L}.$$

The solution is given by

$$(12)\qquad \gamma(z)=\frac{z(1-\delta z)}{R(z)B(z)}\left[\frac{\lambda_2(1+\alpha(1-\delta z))B(z)}{(1-\delta z)R(z^{-1})z}\right]_+$$

where

$$R(z)R(z^{-1})=\lambda_1(1-z)(1-\delta z)(1-z^{-1})(1-\delta z^{-1})+\lambda_2,$$
$$g_{SS}(z)=\sigma^2B(z)B(z^{-1}).$$

If the sales series is a mixed autoregressive, moving average process, say

$$S_t=\frac{N(L)}{D(L)}\mathscr{E}_t,$$

where $N(\ )$ and $D(\ )$ have orders $n$ and $d$, respectively, then it is easily seen that

$$(13)\qquad \gamma(z)=\frac{zP(z)}{N(z)R(z)},$$

[7] For convenience we have added the constant $\delta$. Once the solution is obtained we consider the limiting case when $\delta$ tends to unity (from below).

where $P(\ )$ is of order max $(n, d, 0)$. If in the structural equation (1) $x_t^*$ is the planned level of production for time period $t$, then as in the previous examples the estimating equation is a rational lag distribution.[8]

If the firm attempts to minimize $V$ rather than expected cost in period $t$ (or the average cost per period), then because of the form of the objective function the certainty equivalence principle applies. That is, the one may choose future levels of production to minimize $V$ in which all unknown future variables (in this problem the $S_{t+j}$'s) are replaced by their conditional expectations as of time $t-k$. Thus assuming normality or alternatively restricting ourselves to linear forecasting rules, the future sales may be replaced by their least squares forecasts and $P_{t+j}$ chosen to minimize

$$\hat{V}=\sum_{i=0}^{\infty}\rho^{v}[\lambda_1(P_{t+v}-P_{t+v-1})^2+\lambda_2(I_{t+v}-\alpha\hat{S}_{t+v,t-k})]$$

subject to: $P_{t+v}=\hat{S}_{t+v,t-k}+I_{t+v}-I_{t+v-1}$. The solution (for details see [**2**, **3**] will obviously be a distributed lag between production and expected future sales.[9] If these latter variables are expressed as linear combinations of current and past sales, one ends up with a distributed lag between current (and past) production and the current and lagged values of sales.

*California Institute of Technology*

*Manuscript received April, 1976; final revision received January, 1977.*

## REFERENCES

[**1**] CHILDS, GERALD L.: *Unfilled Orders and Inventories*. Amsterdam: North-Holland Publishing Company, 1967.

[**2**] GRETHER, DAVID M.: "Distributed Lags, Prediction, and Signal Extraction," Cowles Foundation Discussion Paper No. 279, Yale University, 1969.

[**3**] GRETHER, DAVID M.: "Two Notes on Distributed Lags, Prediction, and Signal Extraction," Social Science Working Paper No. 123, California Institute of Technology, 1976.

[**4**] HAY, G. A.: "Production, Price, and Inventory Control," *American Economic Review*, 60 (1970), 531–545.

[**5**] HOLT, C. C., F. MODIGLIANI, J. F. MUTH, AND H. SIMON: *Planning Production, Inventories, and Work Force*. Englewood Cliffs, New Jersey: Prentice-Hall, 1960.

[**6**] JORGENSON, D. W.: "Rational Distributed Lag Functions," *Econometrica*, 34 (1966), 135–149.

[**7**] NERLOVE, MARC: "Distributed Lags and Unobserved Components of Economic Time Series," in *Ten Economic Studies in the Tradition of Irving Fisher*, ed. by W. Fellner, et al. New York: John Wiley, 1967, pp. 127–169.

[**8**] ———: "Lags in Economic Behavior," *Econometrica*, 40 (1972), 221–251.

[**9**] THEIL, H.: *Optimal Decision Rules for Government and Industry*. Amsterdam: North-Holland Publishing Company, 1964.

[**10**] WHITTLE, P.: *Prediction and Regulation by Linear Least Squares Methods*. London: The English Universities Press, 1963.

[8] Other examples are given by [**8**].

[9] [**3**] is available from the author on request.

# LINKED CITATIONS



*You have printed the following article:*

**A Note on Distributed Lags, Prediction, and Signal Extraction**
David M. Grether
*Econometrica*, Vol. 45, No. 7. (Oct., 1977), pp. 1729-1734.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28197710%2945%3A7%3C1729%3AANODLP%3E2.0.CO%3B2-5

---



## [Footnotes]

[6] **Production, Price, and Inventory Theory**
George A. Hay
*The American Economic Review*, Vol. 60, No. 4. (Sep., 1970), pp. 531-545.
Stable URL:
http://links.jstor.org/sici?sici=0002-8282%28197009%2960%3A4%3C531%3APPAIT%3E2.0.CO%3B2-C

[8] **Lags in Economic Behavior**
Marc Nerlove
*Econometrica*, Vol. 40, No. 2. (Mar., 1972), pp. 221-251.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28197203%2940%3A2%3C221%3ALIEB%3E2.0.CO%3B2-R

## References

[4] **Production, Price, and Inventory Theory**
George A. Hay
*The American Economic Review*, Vol. 60, No. 4. (Sep., 1970), pp. 531-545.
Stable URL:
http://links.jstor.org/sici?sici=0002-8282%28197009%2960%3A4%3C531%3APPAIT%3E2.0.CO%3B2-C

**NOTE:** *The reference numbering from the original has been maintained in this citation list.*

# LINKED CITATIONS



[6] **Rational Distributed Lag Functions**
Dale W. Jorgenson
*Econometrica*, Vol. 34, No. 1. (Jan., 1966), pp. 135-149.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28196601%2934%3A1%3C135%3ARDLF%3E2.0.CO%3B2-6

[8] **Lags in Economic Behavior**
Marc Nerlove
*Econometrica*, Vol. 40, No. 2. (Mar., 1972), pp. 221-251.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28197203%2940%3A2%3C221%3ALIEB%3E2.0.CO%3B2-R